Exploring the Perception of Quadruped Motion Retargeting

George Fletcher
The Univeristy of Bath

Yiguo Qiao
The Univeristy of Bath

Rebecca Fribourg
Trinity College Dublin

Jake Deane
The Univeristy of Bath

Rachel McDonnell
Trinity College Dublin

Darren Cosker
The Univeristy of Bath

KEYWORDS

Motion Capture and Retargeting, Datasets, Animal Locomotion, Perception of Animation, Neural Motion Retargeting

ACM Reference Format:
George Fletcher, Yiguo Qiao, Rebecca Fribourg, Jake Deane, Rachel McDonnell, and Darren Cosker. 2021. Exploring the Perception of Quadruped Motion Retargeting. In *Proceedings of MIG '21: Motion, Interaction and Games (MIG'21).* ACM, New York, NY, USA, 2 pages.

1 OVERVIEW

Computer generated quadruped characters play key roles in entertainment - but are expensive to create and difficult to motion capture. So what do we do if we want multiple high quality quadruped characters but only have suitably expressive data for one? Motion retargeting is a solution, but from biomechanics we know that animal movement is based on the minimisation of energy, i.e. an animal will select the most energy efficient gait pattern when moving. Given differences in physiology, size and shape animals will therefore move differently [Hermanson 2004]. If we then decide to perform retargeting a question arises: do people notice if the movement from one quadruped is transferred onto another? This has implications in animation – and questions relating to the Uncanny Valley [Mori et al. 2012]. To start to answer these questions, we substantially build upon an existing motion capture library of dogs [Kearney et al. 2020] and implement two state-of-the-art motion retargeting methods [Aberman et al. 2020; Villegas et al. 2018] on this data, and then conduct a perceptual study with regards to the 'naturalness' of the produced motions. We found that participants gave similar, yet interestingly low, naturalness ratings to both retargeted and original motions.

2 DATA AND RETARGETING METHODS

For animals there is significantly less motion capture data available – mainly due to the practical issues. Likely as a result, work on visual perception of quadrupeds is very limited [Skrba and O'Sullivan 2009]. We added 11 more breeds and several new motion types to [Kearney et al. 2020]. The captured sequences cover seven motions: walking, trotting, jumping, maneuvering over poles, and jumping on and off tables, adding on average 2400 frames per dog at 60fps. We used a 20 camera Vicon motion capture system and the Shogun software to perform capture, solving the same topology skeleton

MIG'21, November 10–12, 2021, Lausanne, Switzerland

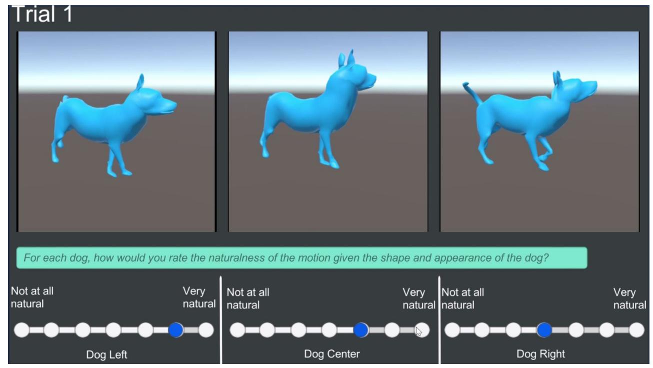

Figure 1: Screenshot of the experimental application developed in Unity 3D.

hierarchy to the marker data to ensure consistency with [Kearney et al. 2020]. Given our motion capture data, we now consider retargeting methods with which to transfer motions between different dogs. For this task, we selected the methods of [Villegas et al. 2018](NKN) and [Aberman et al. 2020](AB) as they are both data-driven unsupervised methods with available implementations online which demonstrate state of the art performance without needing animator expertise. For both we trained them on the entire dataset for performance (as the aim is to share all captured motions between characters) and found the default hyper-parameters to perform best.

3 EXPERIMENTAL DESIGN

In our experiment, we selected three different movement types – *walk*, *trot*(*run*), *pole* – and three different dogs D_1, D_2, D_3, selected to be as physiologically different as possible having average bone lengths of 10.5cm, 8.4cm and 6.6cm respectively. For each movement type, we created animations covering each retargeting combination and each retargeting method.

In our design, we present participants with three animations at once – looping each sequence 3 times resulting in a visualisation that lasts on average 8 seconds. The animations are shown in a row in random order, and consist of a tuple of the *Original* motion for D_i, along with two retargeted motions (AB/NKN) to D_i from the same input clip (with different timing). We visualise the dogs as skinned meshes using linear blend skinning. Figure 1 illustrates an example of this presentation as viewed by a participant taking part in the study. This presentation format allows us to compare how well different algorithms perform when given the same retargeting task under the same conditions. Varying the input breed then allows us to compare whether differences in physiology affects algorithm performance. When presenting a tuple to a participant we ask them to score each animation on a 7-point likert scale against the following question: *"How natural would you rate the motion given the shape and breed of the dog?"*. Note that we do not present viewers with exemplars of how a specific dog would move beforehand to simulate the experience of viewing an animated animal in a video game or movie context, where the only prior is your expectations of how that animal should move. Showing 3 clips, as opposed to one at a time, was chosen to assist the viewer in considering relative naturalness ratings as a natural clip would always be present. Note that we felt removal of the floor plane was necessary due to the possibility of minor floor clipping potentially distracting the viewer. In total, participants had to rate 54 clips (45 unique) of animation and the experiment lasted approximately 10 minutes.

4 RESULTS

Our experiment was carried out by 20 participants in total – 14 males and 6 females – with a mean age of 25.8. A three-way Analysis of Variance (ANOVA) was performed in this analysis. The normality assumption was tested using a Shapiro-Wilk test and when not verified, an Aligned Rank Transformation (ART) was applied on the data. Tukey's post-hoc tests ($\alpha = .05$) were conducted to check significance for pairwise comparisons. In addition, post-hoc tests were corrected using Bonferroni correction. A three-way ANOVA was performed comparing ratings of Naturalness, considering the within-subject factors Method and Type of Movement. We also consider the factor Transition from one dog to another in order to assess the impact of the dog breed (and potentially the dog size) on the results. This analysis was only conducted comparing clips with method *AB* or *NKN* without involving *Original* clips. The results did not highlight any main effect or interaction from our factors, although a tendency appeared for an interaction between the Type of Movement and the Method (p=0.08). Because we were also interested in exploring how participants would differentiate animations from retargeting methods and original motions, we conducted another three-way ANOVA considering the within-subject factors Method (AB, NKN, $Original$), Type of Movement ($walk$, $trot$, $pole$) and Dog (D_1, D_2,D_3). As before, no main effect was found, except a tendency of interaction effect between Dog and Type. As mentioned in the overview, the naturalness ratings of the original clips were relatively low (see Figure 2), which is discussed in the following paragraph.

5 DISCUSSION

State-of-the-art unsupervised neural retargeting techniques are typically trained with bipedal data, thus we have demonstrated the applicability of such methods to quadrupedal data. We are the first to analyse the perception of neural quadruped retargeting and believe that future motion retargeting methods could produce fruitful perceptual studies. Regarding our perceptual study, the results suggest that participants were unable to distinguish between original movements and the retargeting methods when rating the naturalness of the motion. This result is interesting and somewhat surprising, as we expected participants to be able to recognise natural motion when shown retargeted motions alongside, since we felt we could recognise the differences. While this may support the

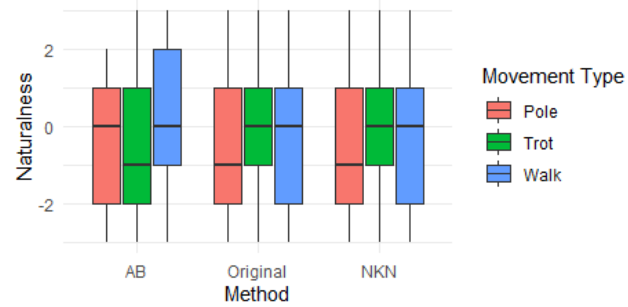

Figure 2: Naturalness ratings depending on Method and Movement Type.

efficiency of the two retargeting methods, we must note that overall naturalness ratings were relatively low, including original motions. We think there could be several reasons for this phenomenon. First of all, this study is based on the very subjective judgement of people and their own interpretation of 'naturalness'. Secondly, the natural dog movements often appear unnatural at times for some clips, especially the neck movement, which may even result in retargeted results leading to more natural looking visual perception (see AB on 'Walk' in Figure 2). We also found the AB method would sometimes produce root rotation artefacts which required removal to produce suitable stimuli. More data per dog would likely improve this.

In future work, it would be interesting to increase the length and exposure time to the clips and explore its impact on the perceived naturalness. Importantly though, our work highlights that natural motions may not appear very natural, thus we invite future work to investigate this effect and its relation to the presentation of motion and the subjective interpretation of 'naturalness'.

REFERENCES

Kfir Aberman, Peizhuo Li, Dani Lischinski, Olga Sorkine-Hornung, Daniel Cohen-Or, and Baoquan Chen. 2020. Skeleton-Aware Networks for Deep Motion Retargeting. *ACM Transactions on Graphics (TOG)* 39, 4 (2020), 62.

John W. Hermanson. 2004. Alexander, R. McNeill. 2003. Principles of Animal Locomotion. Princeton University Press, Princeton, New Jersey, and Woodstock, Oxfordshire, United Kingdom, 371 pp. ISBN 0-691-08678-8. *Journal of Mammalogy* 85, 3 (06 2004), 583–584. https://doi.org/10.1644/1383962

Sinead Kearney, Wenbin Li, Martin Parsons, Kwang In Kim, and Darren Cosker. 2020. RGBD-Dog: Predicting Canine Pose from RGBD Sensors. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).* http://arxiv.org/abs/2004.07788 arXiv: 2004.07788.

Masahiro Mori, Karl F. MacDorman, and Norri Kageki. 2012. The Uncanny Valley [From the Field]. *IEEE Robotics Automation Magazine* 19, 2 (2012), 98–100. https://doi.org/10.1109/MRA.2012.2192811

Ljiljana Skrba and C. O'Sullivan. 2009. Human perception of quadruped motion. In *APGV '09*.

Ruben Villegas, Jimei Yang, Duygu Ceylan, and Honglak Lee. 2018. Neural kinematic networks for unsupervised motion retargetting. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 8639–8648.